Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Curl-free ray mapping in three dimensions for freeform illumination design

Caleb Gannon
Rongguang Liang

SPIE.

Caleb Gannon, Rongguang Liang, "Curl-free ray mapping in three dimensions for freeform illumination design," *Opt. Eng.* **58**(2), 025102 (2019), doi: 10.1117/1.OE.58.2.025102.

Curl-free ray mapping in three dimensions for freeform illumination design

Caleb Gannon and Rongguang Liang*
University of Arizona, College of Optical Sciences, Tucson, Arizona, United States

Abstract. An illumination design method that incorporates the entire 3-D geometry between the optical surface and target plane into the ray mapping calculation has been developed. The resulting mapping relationship is integrable along the specific contours of the optical surface instead of a 2-D plane. The method is capable of creating extremely high performance, off-axis designs (97% uniformity at 54-deg off-axis), and since it explicitly accounts for system geometry, appears to be easily generalizable. [DOI: 10.1117/1.OE.58.2.025102]

Keywords: freeform; illumination design; nonimaging optics.

Paper 181431 received Oct. 4, 2018; accepted for publication Jan. 18, 2019; published online Feb. 8, 2019.

1 Introduction

With ever-growing manufacturing capabilities, freeform surfaces are quickly making their way into the world of optics. In the context of illumination design, these freeform surfaces can provide arbitrary degrees of freedom, allowing for complete control over a given point source light distribution. This freedom comes at the cost of increasing design complexity, which has not been completely solved at the time of this publication. To realize the full capabilities of this freedom, a robust method for generating a freeform optical surface for any conceivable geometry is still needed.

In general, the illumination design process requires the calculation of two parameters: a mapping relationship between the source and target energy distributions and a surface capable of redirecting the light according to that mapping. Although powerful simultaneous solutions for both parameters do exist, for example,[1,2] they are often limited by their complexity making a general solution for a variety of conditions difficult. To overcome this problem, the mapping relationship and surface profile are often calculated in separate steps by first establishing a mapping correspondence between the source and target energy distributions and then constructing a surface normal field that directs the source rays toward the target locations. Finally, an optical surface can be built according to this normal field.

In our previous paper, we pointed out that a lack of geometrical information in the mapping calculation poses a fundamental problem with this two step method.[3] While our previous approach showed a significant improvement over traditional means, it was unable to completely remove the curl in the remaining vector field. This meant that constructing a smooth surface to completely enforce the desired mapping relationship was impossible.[4] At the time of our publication, we believed that this shortcoming was due to the fact that optimum transport theory could only account for the location, not the direction of the light rays. While this is true, it is only a side-effect of the fact that our ray mapping relationship, $\mathbf{u}$, was only constructed for two dimensions:

$$\mathbf{u}:\mathbf{R}^2 \rightarrow \mathbf{R}^2. \tag{1}$$

Although we calculated the light energy distribution in three dimensions, the actual mapping correspondence was performed between x and y values on the lens and target plane. For surfaces with gradual changes in z, this simplification does not pose much of a problem because the cost function of the mapping calculation is mostly driven by the larger differences in x and y. But for surfaces with large height variations (which are especially common when designing for off-axis illumination), the errors from this 2-D approximation become increasingly apparent. See, for example, Fig. 1 where we consider two rays in our mapping calculation. Here, P denotes the surface of the lens and T is the target plane.

While $\vec{a}$ and $\vec{b}$ are clearly oriented in different directions with unequal magnitudes, they travel equal distances along x. If we assume a standard cost functional for our mapping relationship that depends on the distance in x only, then

$$c(\mathbf{P}_x, \mathbf{T}_x) = |\mathbf{P}_x - \mathbf{T}_x|^n, \tag{2}$$

where $\mathbf{P}_x, \mathbf{T}_x$ denote the x co-ordinate on the surface and target, respectively, and $n \in \mathbf{R}^+$ (typically, $n = 2$ to denote a quadratic cost), our calculation will be the same for both vectors. This inhibits our mapping calculation from understanding the full information of the geometry, which leads to mapping relationships, which are integrable along a flat plane but not along the specific shape of the optical surface.

2 Vector Normalization

As an example of this effect, we looked at the design of an off-axis, reflective system with geometry shown in Fig. 2. A ± 40 deg circular aperture collected a Lambertian point source and reflected the light onto a 200 mm $\times$ 200 mm uniform square target 54 deg off-axis.

To evaluate system performance, we computed the residual z component of the curl in the mapping relationship in two ways. First, we looked at the vectors between the x and y components at the input and calculated output

*Address all correspondence to Rongguang Liang, E-mail: rliang@optics.arizona.edu

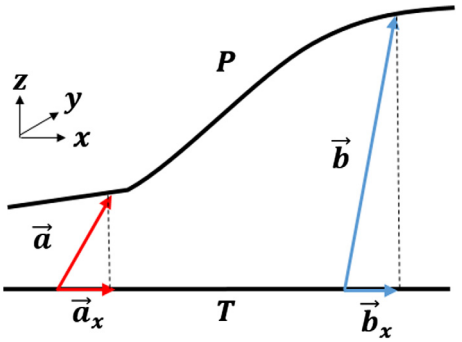

Fig. 1 Ray diagram showing how differently oriented vectors from the target plane **T** to the lens surface **P** can have similar values when projected to the *x*-*y* plane.

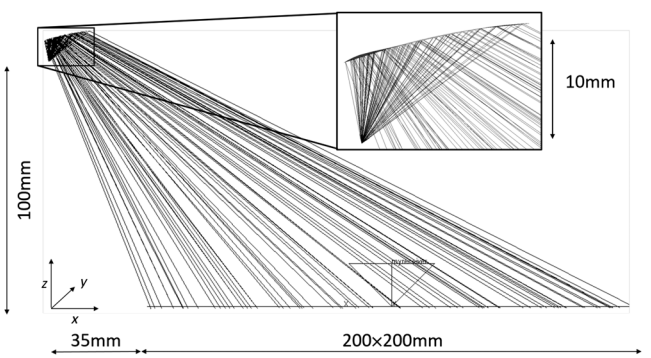

Fig. 2 Test geometry for the 54 deg off-axis design shown in LightTools simulation software. The surface was placed 10 mm above the source and tilted slightly toward the target plane to avoid obscuration from the source. The angular offset was measured as the angle between the surface normal at the center of the target plane and the source.

locations of the mapping relationship in Fig. 3(a). Unsurprisingly, the curl of this vector field was extremely close to zero, as shown in Fig. 3(b), because our mapping calculation was performed on two planar regions between the input and target. But when we then looked at the x and y components of the full 3-D vector field between the surface and target in Fig. 3(c) by normalizing according to

$$\mathbf{u} = \mathbf{T} \to \mathbf{P} = \frac{\left\langle \mathbf{P}_x - \mathbf{T}_x, \mathbf{P}_y - \mathbf{T}_y \right\rangle}{\left| \left\langle \mathbf{P}_x - \mathbf{T}_x, \mathbf{P}_y - \mathbf{T}_y, \mathbf{P}_z - \mathbf{T}_z \right\rangle \right|}, \tag{3}$$

we find there is still curl in the mapping relationship that is not accounted for in the planar approximation. We can see in Fig. 3(d) that this remaining curl corresponds directly to the errors in the final irradiance pattern of our design, which can be seen in Fig. 4.

Since this curl cannot be accounted for using the planar approximation, a method to incorporate the 3-D vector information into the calculation is imperative. However, with a closer inspection of Figs. 3(c) and 3(d), we realize that we have discovered exactly such a method. By normalizing the vectors, the x and y components are scaled appropriately so that the distance traveled in z is implicitly embedded into the calculation of vector curl. Effectively what we have done by normalizing the x and y components is to translate the problem of point-to-point mapping into a new problem, where points are mapped to directions, which are aimed at the target locations. Figure 5 shows this more explicitly.

In this paper, we propose this point-to-wavefront mapping through vector normalization as a way to perform extremely accurate 3-D mapping calculations within the framework that has been developed for 2-D optimal transport.

3 Reducing Curl for Optimal Transport

The L^2 Monge–Kantorovich problem has applications reaching far beyond optics and into fields such as statistics, functional analysis, and atmospheric sciences.[5,6] An important result from these studies is the existence of a unique optimal mapping, $\tilde{\mathbf{u}}$, which can be described as the gradient of a convex function w, i.e., $\tilde{\mathbf{u}} = \nabla w$.[7–12]

To arrive at this optimal solution, since our original mapping $\mathbf{u}$ is a vector field, we can always decompose it into two terms:

$$\mathbf{u} = \nabla w + \boldsymbol{\chi}, \tag{4}$$

where we select ∇w as the curl-free term, using the property that for any scalar potential, the curl of the gradient is always zero, i.e., $\nabla \times \nabla w = 0$. Subsequently, the second term then contains all of the curl and is divergence-free, such that $\mathrm{div}\boldsymbol{\chi} = 0$. Importantly, we can see that to reach the optimal mapping $\tilde{\mathbf{u}}$, we must reduce the term containing all of the curl, $\boldsymbol{\chi}$, to zero.

To do this, we utilize the evolution equation proposed by Haker et al.:[7]

$$\frac{d\mathbf{u}}{dt} = -\frac{1}{\mu_0} D\mathbf{u} \nabla^{\perp} (\Delta^{-1} \mathrm{div}\mathbf{u}^{\perp}), \tag{5}$$

where $\Delta^{-1}\mathrm{div}\mathbf{u}^{\perp}$ denotes the solution of the Poisson equation:

$$\Delta f = -\mathrm{div}\mathbf{u}^{\perp}, \tag{6}$$

where μ_0 is the density map in the source domain, $D\mathbf{u}$ is the Jacobian determinant, and $(x, y)^{\perp}$ denotes a rotation by 90 deg in $\mathbf{R}^2$. Given an initial mapping relationship, $\mathbf{u}$, this equation solves for a new mapping, $\tilde{\mathbf{u}}$, by removing the curl through a standard gradient descent evolution.

Previously, optimal transport problems of this type have been solved between parallel planes in $\mathbf{R}^2$.[13] With the proposed method, we are now able to account for the entire 3-D geometry of the problem by mapping to the normalized wavefront of a vector field by calculating $\mathbf{u}$ according to Eq. (3).

One added difficulty of working with normalized vectors is that during the iterative process, the vector components u_x and u_y point toward varying locations on the surface. This means that with each iteration, the vectors need to be renormalized according to the new surface positions. Since the changes are inherently small, we can incorporate a modified version of Euler's method to solve for the intersection point between the vector and the tangent to the surface normal, as shown in Fig. 6.

Knowing that a vector starts from $\mathbf{P}_1$ in the direction of $\vec{v}_1$ and the tangent vector starts from $\mathbf{P}_2$ in the direction of $\vec{v}_2$, we can create a system of parametric equations for the lines as a function of arbitrary length variables t_1 and t_2, which can

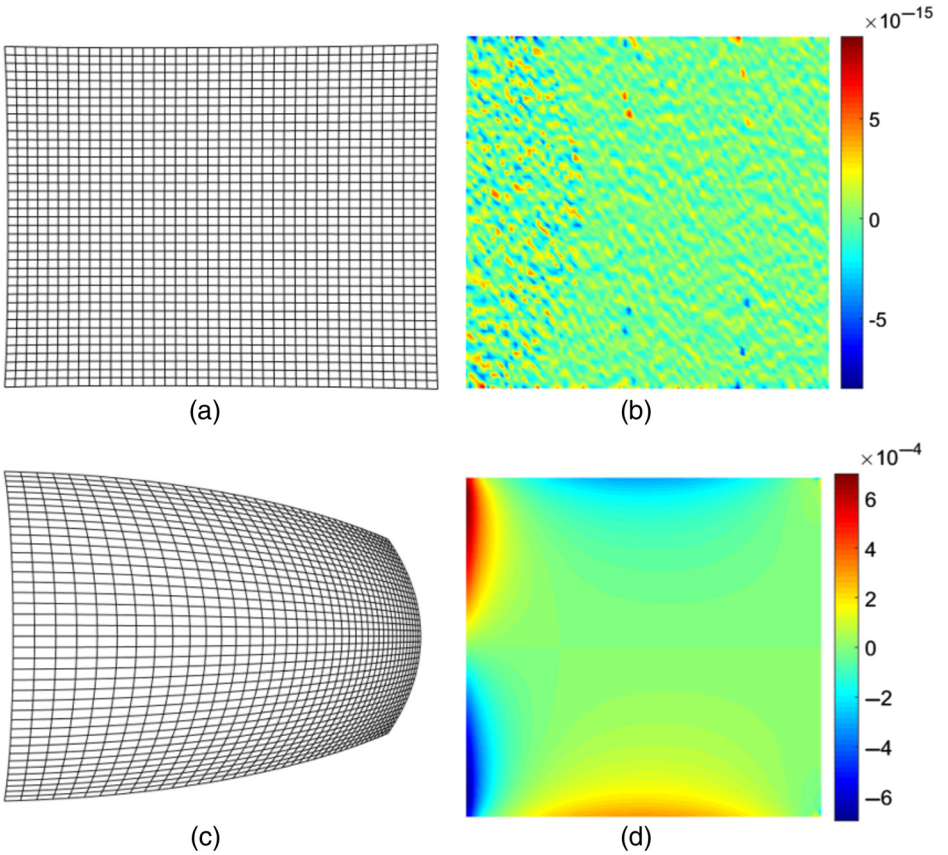

Fig. 3 (a) x-y components of the mapping relationship demonstrated as the deformation of a unit grid. (b) Z component of the curl in the mapping relationship between the calculated x and y values in the source and target plane (note the error is on the order of 10^{-15}). (c) The x and y components of the normalized mapping relationship. (d) Z component of the curl in the mapping relationship using the x and y components of the normalized 3-D vectors between the lens surface and the target.

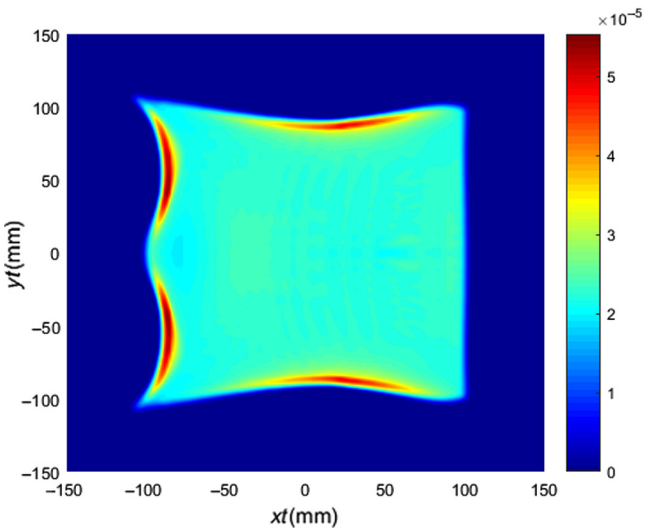

Fig. 4 Resulting irradiance distribution from the surface constructed according to the 2-D mapping relationship analyzed in Fig. 3.

be solved simultaneously for the new co-ordinate on the surface where the two intersect as follows:

$$\mathbf{P}_1 - \mathbf{P}_2 = t_2\vec{v}_2 - t_1\vec{v}_1. \tag{7}$$

Using the new surface locations, $\mathbf{P}' = t_1\vec{v}_1$, we can then recalculate the normalized vector field according to Eq. (3).

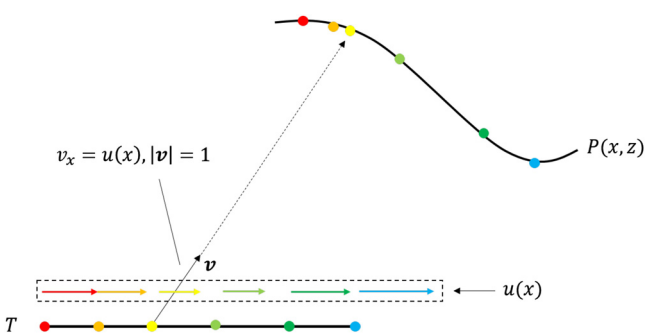

Fig. 5 By normalizing the x components of the mapping $u(x)$, any vector of unit magnitude with x component equal to $u(x)$ will land on the desired surface location. This method works in the y direction as well but was shown in 2-D for simplicity.

4 Jacobian of the Warping Map

One important aspect of this approach to address is the fact that we are calculating transport in a distorted frame. Although normalization works with vector fields, when dealing with physical properties, such as the light distribution in space, it does not necessarily make sense. In Fig. 7, we can see the difference between the x-y locations on the lens surface (a) and the x-y values of the normalized vector field (b).

Typically, distortions between the source and target geometries are accounted for using the Jacobian determinant.

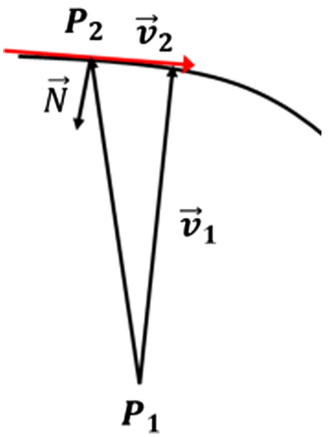

Fig. 6 Given a small enough step size, the new surface point can be approximated as the intersection between the calculated vector $\vec{v}_1$ and the tangent vector to the surface normal $\vec{v}_2$.

Take, for example, Fig. 8, where a nonlinear map f turns small squares into distorted parallelograms.

While the differential area of the squares on the left is easily calculated as the multiple of each differential, $dA = dxdy$, the warped geometry requires an additional factor corresponding to the nature of f such that

$$dA = D\mathbf{u}dxdy, \tag{8}$$

where $D\mathbf{u}$ is referred to as the Jacobian determinant.

Since the Jacobian works to conserve differential areas between the source and target planes, the x-y locations on the actual surface need to be used when calculating the Jacobian to produce a mass-preserving mapping relationship. However, if the stable stepsize described in Ref. 7 is to be used, a uniform scaling needs to be introduced such that the total magnitude of the Jacobian calculated on the optical surface is equal to that when calculated from the vector field. Otherwise, a separate scheme for determining stepsize will need to be developed.

In the following sections, we demonstrate the effectiveness of the proposed method in designing off-axis illumination systems, one of the most challenging problems in illumination design.

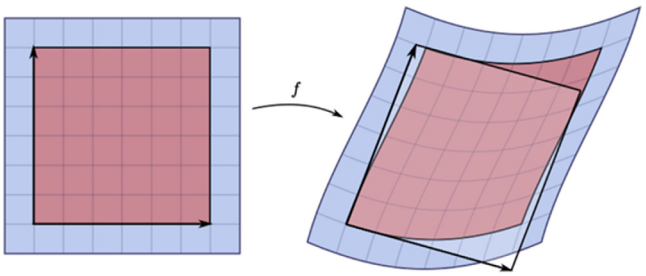

Fig. 8 A nonlinear map warps the shape of differential areas from squares to parallelograms. Image from Ref. 14.

5 Reflective Freeform Surface

Utilizing the same geometry described in Fig. 2, we calculated a mapping relationship using the method described in Ref. 3. Then, we utilized the procedure described in Sec. 3 to remove the remaining curl in the 3-D vector field. Equation (6) was solved using MATLABS PDE toolbox, and the Jacobian of the warping was evaluated with a standard upwind scheme. The step size we used for iterations was the one described by Haker et al.:[7]

$$\min_{x,i}\left|\frac{1}{\mu_0}(\nabla^{\perp}\Delta^{-1}\mathrm{div}\mathbf{u}^{\perp})_i\right|^{-1} \tag{9}$$

to ensure stability during calculation. The results after 228 iterations are shown below in Fig. 9 with the convergence plot in Fig. 10. As you can see, the peak curl was reduced by more than an order of magnitude.

After building an optical surface to enforce the calculated mapping in the same manner as in Refs. 3 and 15, the performance of this method was tested against our previous method below in Fig. 11. The uniformity within the target region is quite good at 97% even at 54 off-axis.

6 Refractive Freeform Surface

To demonstrate the versatility of this method, we also designed a 20-mm off-axis, refractive lens using a simple x-y mapping method to generate our initial surface (see Secs. 8 and 9 for details). A square subtending a 71 collection angle of a Lambertian point-source was directed onto a 200 mm × 200 mm target 100 mm away. The first surface

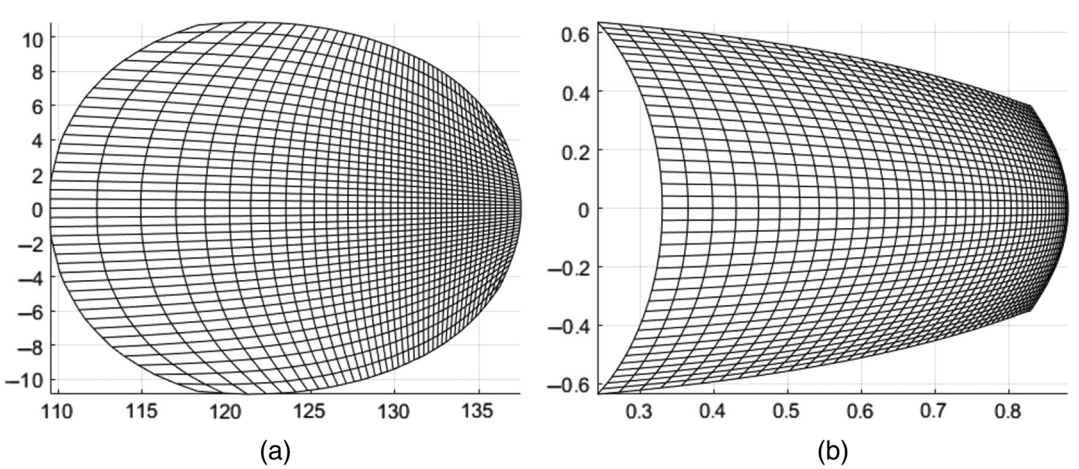

Fig. 7 x-y components of (a) the optical surface and (b) the normalized vector field (down sampled to prevent aliasing).

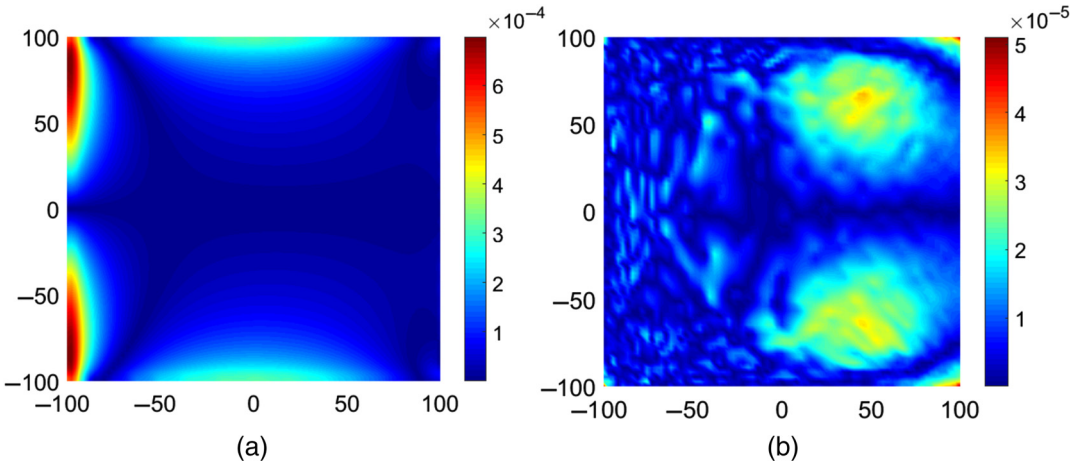

Fig. 9 Residual curl of the mapping using (a) our previous method and (b) the proposed method.

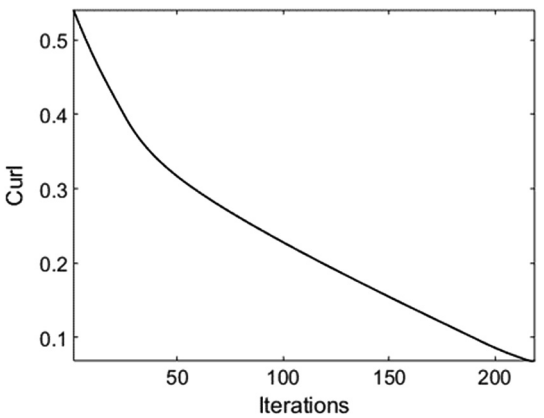

Fig. 10 Convergence plot of the total curl versus number of iterations. The total runtime for this example was just over 2 min.

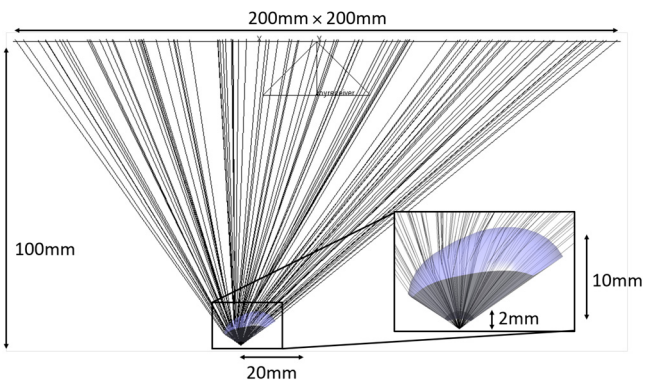

Fig. 12 Test geometry for the off-axis refractive design shown in LightTools simulation software.

was a sphere of radius 2 mm while the second surface was freeform with a height at the vertex of 10 mm. A diagram of this setup is shown in Fig. 12.

After using Eq. (3) to calculate the vector field between the surface and target plane, we used the method in Sec. 3 to remove the remaining curl. The results after 69 iterations are shown below in Fig. 13 with the convergence plot in Fig. 14. A careful inspection of the remaining curl reveals a direct correspondence to the triangular mesh generated to solve the PDE.

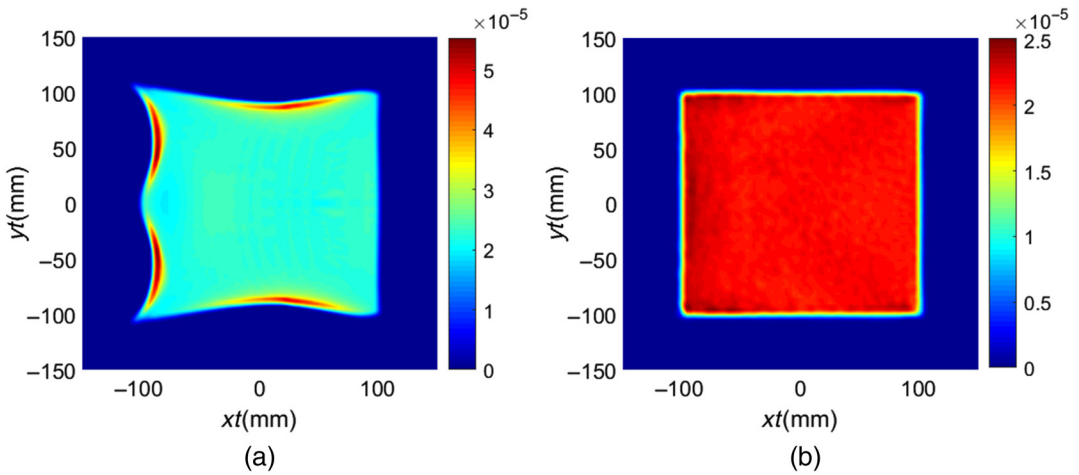

Fig. 11 Irradiance patterns at 54 off-axis for (a) our previous method and (b) the proposed method. The corresponding uniformities within the 200 mm × 200 mm target are 56% and 97%, respectively. The irradiance uniformity was calculated as 100%(1-RMS deviation/mean) on the entire 101 × 101 grid with a 3 pixel smoothing kernel to reduce statistical error from the raytracing.

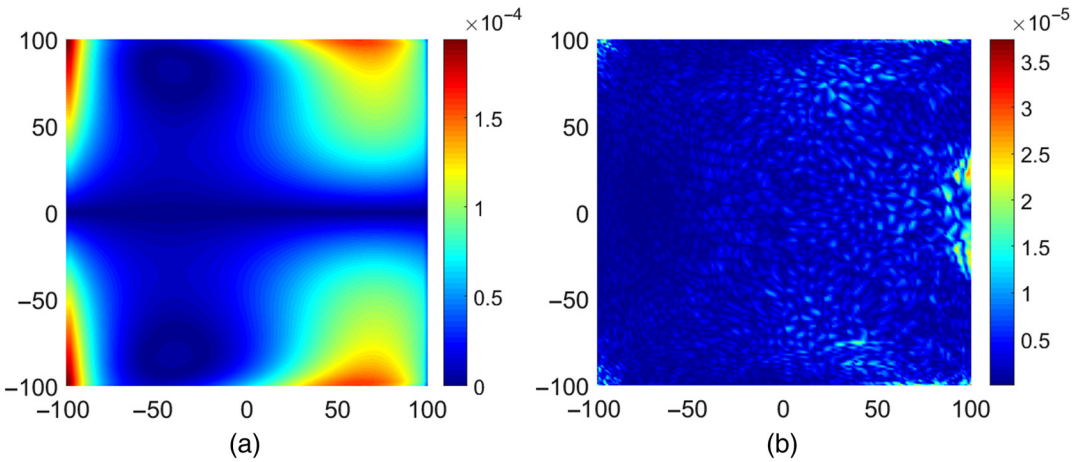

Fig. 13 Residual curl of the mapping from (a) the *X*-*Y* mapping method and (b) after curl reduction.

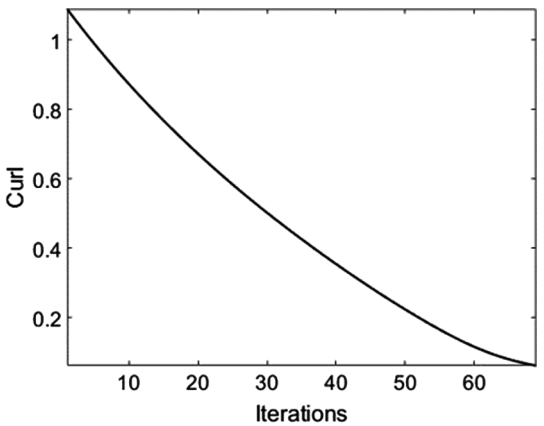

Fig. 14 Convergence plot of the total curl versus number of iterations. The total runtime for this example was just over a minute.

We constructed an optical surface using the same method described in Sec. 5, with the addition of a spherical surface concentric to the source, which was generated by setting the input and output vector fields equal. The results before and after curl reduction are shown below in Fig. 15.

Although the attainable performance of this method is slightly worse when starting with a higher initial curl from the *x*-*y* mapping, this example demonstrates the generalizability of the proposed method as an easy way to improve any mapping calculation whether it is a far off-axis reflector requiring optimal transport calculations or a quick design using simple separation and integration.

7 Conclusion

The above method has been shown to construct an extremely uniform illumination system even in a difficult setup such as off-axis design. Additionally, since the method is only a slight modification to the already existing optimum transport theory, we believe it should easily be generalizable and are optimistic that it will work in any framework in which a wavefront of 3-D vectors can be substituted for mapping co-ordinates, where potential applications extend beyond the scope of illumination design.

Currently, the key limitation in this method is that we are still working with the *x*-*y* co-ordinates of what is a 2-D surface embedded in 3-D space. In regions where the surface doubles over itself (as is typical in hemispherical refractive

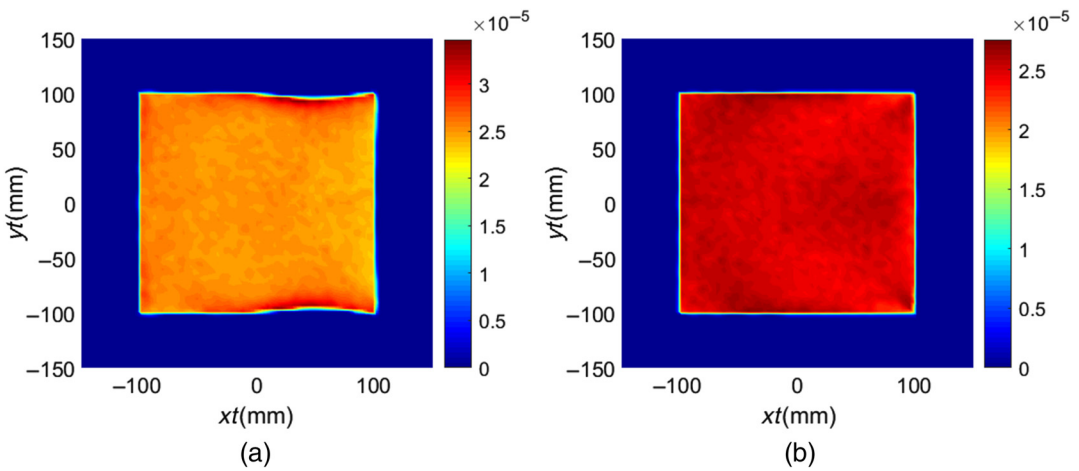

Fig. 15 Irradiance patterns generated by (a) the *x*-*y* mapping calculated on a dummy plane and (b) the proposed method. The corresponding uniformities within the 200 mm × 200 mm target are 84% and 97%, respectively. The irradiance uniformity was calculated as 100%(1-RMS deviation/mean) on the entire 101 × 101 grid with a 3 pixel smoothing kernel to reduce statistical error from the raytracing.

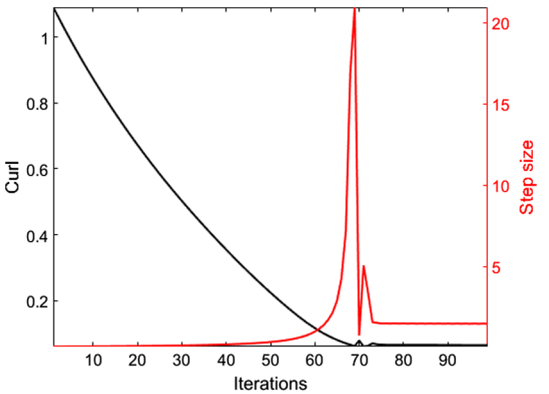

Fig. 16 Convergence plot of the total curl versus number of iterations and the step size used for each iteration.

lenses) or the gradient of the x-y locations differs from the surface gradient [see, for example, the residual light variations in Fig. 15(b)], the method is not fully able to account for these variations. To overcome this, we need to implement the calculations of Secs. 3 and 4 [in particular Eq. (5)] along the surface of the lens, which will be the subject of future work.

8 Appendix A: Convergence and Stability

For the examples described in Secs. 5 and 6, our stopping point for the curl reduction was set to occur after the first time the curl increased from one iteration to the next. Although the original method described in Ref. 7 asymptotically approached a global minimum, in this approach, we are changing the vector field after each iteration. Eventually, we reach a point where the improvement in the vector field during each iteration is of the same magnitude as the change in the vector field introduced by normalization. Looking at Fig. 16, we can see the effect this has on system performance.

After reaching the minimum curl, the variable step size is used in calculation spikes (in response to the tiny value of vector curl). This spike is amplified by the vector normalization and forces the calculation out of its local minimum before settling into a slightly higher region that remains constant between the alternating vector normalization and curl reduction steps. Although this region itself is still quite stable, we found that the best performance is achieved by stopping before this happens.

This value of the minimum point depends on the geometry of the system, the quality of the initial mapping relationship, and the resolution of the PDE solver used. If the initial mapping relationship is too poor, the Jacobian can become numerically unstable. See Fig. 17 for an example.

To solve this problem, a compromise needs to be made between mass preservation and curl reduction. Usually, this is as simple as applying a smoothing function to the Jacobian.

9 Appendix B: X-Y Mapping

To calculate the initial transport map in Sec. 6, we used separation of variables on an intensity distribution sampled on a plane in front of the source. Our solution is almost identical to Sec. 3.2 found in Ref. 7 but is included here for convenience to the reader. According to energy conservation, we know that the total energy between the two regions must be preserved, which we can write as follows:

$$\iint I_1(x,y)\mathrm{d}x\mathrm{d}y = \iint I_0(x,y)\mathrm{d}x\mathrm{d}y. \tag{10}$$

We first establish the mapping along the x axis, $u_x(x)$, that we define by

$$\int_0^{u_x(x)} \int I_1(\chi,y)\mathrm{d}\chi\mathrm{d}y = \int_0^{x} \int I_0(\chi,y)\mathrm{d}\chi\mathrm{d}y. \tag{11}$$

Here, I_0 and I_1 denote the source and target irradiance distributions, and χ is a dummy variable. Differentiating the above with respect to x yields the following:

$$u_x'(x) \int I_1(u_x(x),y)\mathrm{d}y = \int I_0(x,y)\mathrm{d}y. \tag{12}$$

We can then establish a mapping along the y axis, $u_y(x,y)$ according to

$$u_x'(x) \int_0^{u_y(x,y)} I_1(u_x(x),y)\mathrm{d}y = \int_0^{y} I_0(x,y)\mathrm{d}y, \tag{13}$$

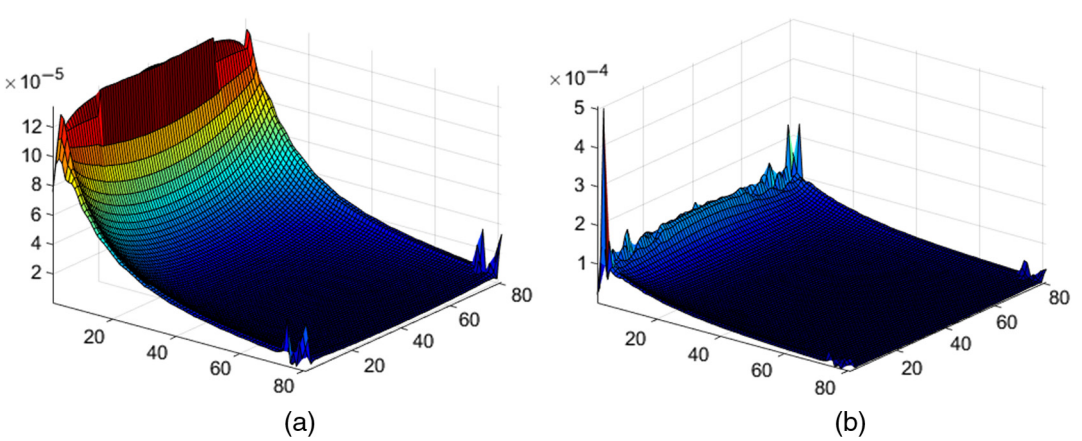

Fig. 17 Jacobian of the warping map (a) at the start of the iterations and (b) when the iterations become unstable. On a small gridsize, those abrupt peaks can lead to numerical instability, hindering performance.

where we introduce another dummy variable, η. Differentiating the above with respect to y gives us as follows:

$$u_x'(x)u_y'(x,y)I_1[u_x(x),u_y(x,y)] = I_0(x,y), \tag{14}$$

where we can substitute $D\mathbf{u} = u_x'(x)u_y'(x,y)$ and $\mathbf{u} = (u_x(x),u_y(x,y))$ to arrive at the mass preserving relationship:

$$D\mathbf{u}I_1(\mathbf{u}) = I_0(x,y). \tag{15}$$

Knowing that this mapping satisfies the requirement of mass preservation, we can go back and solve for $u_x(x)$ and $u_y(x,y)$. In our example, where I_1 is a uniform rectangle, we arrive at the following mapping relationship:

$$u_x(x_i) = \frac{\int_0^{x_i}\int_0^Y I_0(x,y)D\mathbf{u}\mathrm{d}x\mathrm{d}y}{\int_0^X\int_0^Y I_0(x,y)D\mathbf{u}\mathrm{d}x\mathrm{d}y}, \tag{16}$$

$$u_y(x_i,y_{i,j}) = \frac{\int_0^{y_{i,j}} I_0(x_i,y)D\mathbf{u}\mathrm{d}y}{\int_0^Y I_0(x_i,y)D\mathbf{u}\mathrm{d}y}, \tag{17}$$

where X and Y denote the size of the sampling plane.

Acknowledgments

This research was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health award number UH2EB022623 and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1746060.

References

1. R. De Leo, C. E. Gutiérrez, and H. Mawi, "On the numerical solution of the far field refractor problem," *Nonlinear Anal.* **157**, 123–145 (2017).
2. R. Wu et al., "Freeform illumination design: a nonlinear boundary problem for the elliptic Monge–Ampére equation," *Opt. Lett.* **38**(2), 229–231 (2013).
3. C. Gannon and R. Liang, "Ray mapping with surface information for freeform illumination design," *Opt. Express* **25**(8), 9426–9434 (2017).
4. F. R. Fournier, W. J. Cassarly, and J. P. Rolland, "Fast freeform reflector generation using source-target maps," *Opt. Express* **18**(5), 5295–5304 (2010).
5. J.-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem," *Numerische Math.* **84**(3), 375–393 (2000).
6. M. J. Cullen and R. J. Purser, "An extended Lagrangian theory of semi-geostrophic frontogenesis," *J. Atmos. Sci.* **41**(9), 1477–1497 (1984).
7. S. Haker et al., "Optimal mass transport for registration and warping," *Int. J. Comput. Vision* **60**(3), 225–240 (2004).
8. Y. Brenier, "Polar factorization and monotone rearrangement of vector-valued functions," *Commun. Pure Appl. Math.* **44**(4), 375–417 (1991).
9. M. Knott and C. S. Smith, "On the optimal mapping of distributions," *J. Optim. Theory Appl.* **43**(1), 39–49 (1984).
10. W. Gangbo and R. J. McCann, "The geometry of optimal transportation," *Acta Math.* **177**(2), 113–161 (1996).
11. W. Gangbo, "An elementary proof of the polar factorization of vector-valued functions," *Arch. Ration. Mech. Anal.* **128**(4), 381–399 (1994).
12. G. Strang and K. Aarikka, *Introduction to Applied Mathematics*, Vol. **16**, Wellesley-Cambridge Press, Wellesley, Massachusetts (1986).
13. A. Bäuerle et al., "Algorithm for irradiance tailoring using multiple freeform optical surfaces," *Opt. Express* **20**(13), 14477–14485 (2012).
14. Wikimedia, "Jacobian matrix and determinant," https://en.wikipedia.org/w/index.php?curid=46479614 (3 October 2018).
15. Z. Feng, B. D. Froese, and R. Liang, "Freeform illumination optics construction following an optimal transport map," *Appl. Opt.* **55**(16), 4301–4306 (2016).

Caleb Gannon is a PhD student at the University of Arizona College of Optical Sciences. He received his BS degree in engineering physics from Rose–Hulman Institute of Technology in 2015. His current research interests include freeform optical design for imaging and nonimaging systems.

Biography of the other author is not available.